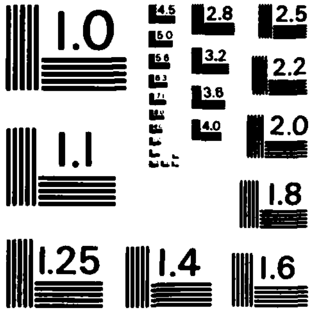
AD-A135 275 ON USING SELECTION PROCEDURES WITH BINOMIAL MODELS(U) 1/1
PURDUE UNIV LAFAYETTE IN DEPT OF STATISTICS
S S GUPTA ET AL. OCT 83 TR-83-44 N00014-75-C-0455
UNCLASSIFIED F/G 12/1 NL

END
DATE
FILMED
1-84
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

12

AD-A135 275

PURDUE UNIVERSITY

DEPARTMENT OF STATISTICS

12

ON USING SELECTION PROCEDURES WITH BINOMIAL MODELS

by

Shanti S. Gupta*
Purdue University

Gary C. McDonald
General Motors Research Laboratories

Technical Report #83-44

Department of Statistics
Purdue University

October 1983

*The research of this author was supported by the Office of Naval Research Contract N00014-75-C-0455 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

ON USING SELECTION PROCEDURES WITH BINOMIAL MODELS

by

Shanti S. Gupta*
Purdue University

Gary C. McDonald
General Motors Research Laboratories

1. Introduction

A common problem that arises in practice is the comparison of k (≥ 2) Bernoulli processes (or populations) $\pi_1,\ldots,\pi_k$ with unknown parameters $p_1,\ldots,p_k$, respectively, where the p_i's denote the 'success' probabilities. For example, the k Bernoulli processes can represent the manufacturing processes of k vendors and p_i can then denote the probability that an item manufactured by the ith vendor will conform to specifications. Let $p_{[1]} \leq \cdots \leq p_{[k]}$ denote the ordered parameters. It is assumed that there is no prior knowledge regarding the correct pairings of the ordered and the unordered p_i's. The vendors (or the processes) are ranked according to the values of p_i's. The vendor associated with $p_{[k]}$, the largest p_i, is called the <u>best</u>.

Let $X_1,X_2,\ldots,X_k$ denote the number of conforming items from these vendors based on a random sample of n items from each manufacturing process. Our interest is to define a statistical procedure based on $X_1,\ldots,X_k$ to select a nonempty subset of the k vendors with a guarantee of minimum probability P* that the best vendor is included in the selected subset. Selection of any subset which includes the best is called a <u>correct selection</u> (CS). Thus the probability of a correct selection using a rule R, $P(CS|R)$, should satisfy the condition that

*The research of this author was supported by the Office of Naval Research Contract N00014-75-C-0455 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

$$P(CS|R) \geq P^* \tag{1.1}$$

whatever be the unknown values of the p_i's. This condition is generally referred to as the P*-condition. Obviously, for a meaningful problem, $1/k < P^* < 1$.

Any procedure R that satisfies (1.1) is a valid procedure. We need to establish a criterion for evaluating the performances of valid procedures. One such criterion is the expected value of S, the number of populations included in the selected subset. S is known as the subset size and it is a positive integer-valued random variable. One may also consider the related quantity E(S'), where S' denotes the number of non-best populations included in the selected subset. Let α_i denote the probability of including the process associated with $p_{[i]}$, $i = 1,\ldots,k$. Obviously, α_k = PCS. It is also easy to see that

$$\begin{cases} E(S) = \alpha_1 + \ldots + \alpha_k \\ E(S') = \alpha_1 + \ldots + \alpha_{k-1}. \end{cases} \tag{1.2}$$

The α_i's are called the individual selection probabilities. One may also consider a criterion which combines E(S) and PCS. Such a criterion, namely, E(S)/PCS has been considered in the literature. All these criteria that are used to evaluate a valid procedure are called the operating characteristics of the procedure. In our present study, we use the expected subset size and the individual selection probabilities. Also, we study these for rules of the same form as R for different values of n and d for a given k. In other words, we are not considering a single value of d corresponding to a given P*.

For an understanding of the importance and need of these decision problems in industry, reference should be made to Deming (1982).

2. The Gupta-Sobel Rule

Gupta and Sobel (1960) proposed and studied a rule R defined as follows.

R: Select π_i if and only if $X_i \geq \max_{1 \leq j \leq k} X_j - d$ where $d = d(k,n,P^*)$ is the smallest nonnegative integer such that the P^*-condition is satisfied. The constant d satisfies

$$(2.1) \qquad \inf_{\Omega} P(CS|R) \geq P^*$$

where $\Omega = \{\underline{p} | \underline{p} = (p_1,\ldots,p_k),\ 0 \leq p_i \leq 1,\ i = 1,\ldots,k\}$ is the parameter space. Gupta and Sobel (1960) have shown that the infimum on the left-hand side of (2.1) is attained when $p_1 = \ldots = p_k$. Thus, we evaluate $P(CS|R)$ for $p_1 = \ldots = p_k = p$ (say) and rewrite (2.1) as

$$(2.2) \qquad \inf_{0 \leq p \leq 1} \sum_{j=0}^{n} \binom{n}{j} p^j (1-p)^{n-j} \left\{ \sum_{y=0}^{d+j} \binom{n}{y} p^y (1-p)^{n-y} \right\}^{k-1} \geq P^*.$$

There is no known result regarding the value of p for which the infimum in (2.2) is attained except in the special case of $k = 2$. When $k = 2$, the infimum is attained for $p = 0.5$. When n is large enough to justify normal approximation, the constant d is approximated by $\tilde{d} = (u\sqrt{n}-1)/2$ which is given by

$$(2.3) \qquad \int_{-\infty}^{\infty} \Phi^{k-1}(x+u)\varphi(x)dx = P^*$$

where Φ and φ denote the cdf and the density of a standard normal variable, respectively. The values of d have been tabulated by Gupta and Sobel (1960) for $k = 2(1)20(5)50$ and $n = 1(1)20(5)50$.

3. Operating Characteristics

Let us assume without loss of generality that $p_1 \leq \cdots \leq p_k$. As we pointed out earlier, we consider the rule: Select π_i if and only if $X_i \geq \max_{1 \leq j \leq k} X_j - d$, where $0 \leq d \leq n$. The operating characteristics studied are the expected subset size and the individual selection probabilities. We consider two types of parametric configurations, namely, (1) the slippage configuration defined by $p=p_1=\cdots=p_{k-1} = p_k-\delta$, $0 < \delta < 1-p$, and (2) the equi-spaced parametric configuration defined by $p_{i+1}-p_i = \delta$, $i = 1,\ldots,k-1$, $0 < \delta < (1-p)/(k-1)$. For convenience, let

$$(3.1)\quad \begin{aligned} b(x;\, n,p) &= \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0,1,\ldots,n \\ B(t;\, n,p) &= \sum_{x=0}^{t} b(x;\, n,p), \quad t = 0,1,\ldots,n. \end{aligned}$$

3.1. Slippage Configurations

For the configuration $(p,p,\ldots,p,p+\delta)$, $0 < \delta < 1-p$, we get

$$(3.2)\quad \left\{ \begin{aligned} PCS \equiv \alpha_k &= \sum_{x=0}^{n} b(x;\, n,p+\delta)[B(x+d;\, n,p)]^{k-1}, \\ \alpha_i &= \sum_{x=0}^{n} b(x;\, n,p)B(x+d;\, n,p+\delta)[B(x+d;\, n,p)]^{k-2}, \\ &\qquad i = 1,\ldots,k-1. \end{aligned} \right.$$

Any specified non-best population has the same probability of being selected and we denote this by P(NCS). Also, $E(S) = (k-1)\alpha_1 + PCS$.

We present tables and graphs for the operating characteristics in the case of three slippage configurations. These are given by the following pairs of p and δ values:

(I) $p = .50$, $\delta = .10$, (II) $p = .75$, $\delta = .05$, (III) $p = .90$, $\delta = .03$. Tables 1A through 1C give the values of PCS, P(NCS), and E(S) for $k = 3,5,10,15$; $d = 2,3,4,5$; and $n = 5(5)50$ in the case of the three configurations I - III. Figure 1 shows the graph of E(S) as a function of n for the rule with $d = 2$ for $k = 3,5,10$ when the slippage configuration is given by $p = .90$ and $\delta = .03$. This figure also shows for $n = 10(10)50$, the value of PCS when $\delta = 0$, that is, when all the parameters are equal to .90. Figures 2A through 2C and 3A through 3C are graphs of E(S) as a function of n for $d = 2,3,4,5$. The first set (2A-2C) is for the slippage configuration with $p = .75$ and $\delta = .05$ when $k = 3,5,10$ and the second set (3A-3C) is for the configuration with $p = .90$ and $\delta = .03$ when $k = 3,5,10$. These results and examples are discussed in the next section.

For sufficiently large n, one can use the normal approximation and obtain

$$
(3.3)\quad \begin{cases} PCS \approx \int_{-\infty}^{\infty} \Phi^{k-1}\left[x\sqrt{\frac{(p+\delta)(q-\delta)}{pq}} + \frac{d+n\delta+1/2}{\sqrt{npq}}\right]\varphi(x)dx, \\ \alpha_i \approx \int_{-\infty}^{\infty} \Phi^{k-2}\left[x + \frac{d+1/2}{\sqrt{npq}}\right]\Phi\left[\sqrt{\frac{pq}{(p+\delta)(q-\delta)}}\, x + \frac{1/2+d-n\delta}{\sqrt{n(p+\delta)(q-\delta)}}\right]\varphi(x)dx, \end{cases}
$$

$$i = 1,\dots,k-1,$$

where $q = 1-p$.

3.2 Equi-spaced Parametric Configuration

For the configuration $(p,p+\delta,\dots,p+(k-1)\delta)$, $0 < \delta < (1-p)/(k-1)$, we have

$$
(3.4)\quad \alpha_i = \sum_{x=0}^{n} b(x;\, n,p+(i-1)\delta) \prod_{j\neq i} B(x+d;\, n,p+(j-1)\delta),\ i = 1,\dots,k.
$$

We note that α_i is the probability of including the non-best population with parameter $p + (i-1)\delta$, $i = 1,\ldots,k-1$, and α_k is the PCS. For large n, the normal approximation yields

$$\text{(3.5)} \qquad \alpha_i \approx \int_{-\infty}^{\infty} \prod_{j \neq i} \Phi\left[\sqrt{\frac{\theta_i(1-\theta_i)}{\theta_j(1-\theta_j)}}\, x + \frac{d+1/2+(i-j)n\delta}{\sqrt{n\theta_j(1-\theta_j)}}\right]\varphi(x)dx,\ i = 1,\ldots,k,$$

where $\theta_i = p+(i-1)\delta$, $i = 1,\ldots,k$.

4. Numerical Illustrations and Discussion

For the purpose of illustrating our rule and the use of the tables, let us assume that we have $k = 5$ processes $\pi_1,\ldots,\pi_5$ with parameters in a slippage configuration $(p,p,p,p,p+\delta)$ with $p = .90$ and $\delta = .03$. Let x_i denote the realized value of X_i, $i = 1,\ldots,5$, in a (hypothetical) data set based on random samples of size $n = 30$ from these processes. Suppose that $x_1 = 27$, $x_2 = 25$, $x_3 = 24$, $x_4 = 22$ and $x_5 = 28$. Consider the rule R with $d = 2$. This rule selects those processes for which $x_i \geq \max\limits_{1 \leq j \leq 5} x_j - 2$; in other words, those for which $x_i \geq 28-2 = 26$. This results in the selection of the subset $\{\pi_1,\pi_5\}$. Note that π_5 is the best process. Consulting Table 1C we find that the PCS for our rule is 0.86. Further, the probability of including a specific non-best population, say, π_3 is 0.66 and the expected subset size is 3.52. If all the processes were identical with $p = .90$, the PCS would have been only 0.702 (see Figure 1).

Given the 5 processes, one may want to decide on a pair (n,d) to use in R. If we can assume that the parameters are in a slippage configuration with, say, $p = .75$ and $\delta = .05$, we can look for a pair (n,d) for which the PCS is at least a specified number, say, .90. If there are more than one such pair with same n, we will naturally take the pair which has the

smallest E(S). Consulting Table 1B, we have the options listed below.

n	d	PCS	E(S)
5	2	.96	4.65
10	3	.96	4.56
15	3	.92	4.13
20	4	.95	4.32
25	4	.93	4.06
30	4	.91	3.82
35	4	.90	3.62
40	5	.93	3.91
45	5	.93	3.74
50	5	.92	3.60

It should be noted that because of the discrete nature of the distribution involved, an increase in n does not necessarily produce a better option. In the range of our table, we see that the best option is n = 50 and d = 5.

Alternatively, one may want to set an upper bound for E(S)/k, the expected proportion of populations selected. If we set this bound as .80, then we look for pairs (n,d) for which E(S) = 5 x .80 = 4. If there are more than one such pair with same n, we take the pair for which the PCS is maximum. Consulting Table 1B again, we have the following options.

n	d	E(S)	PCS
10	2	3.89	.87
15	2	3.37	.81
20	3	3.77	.88
25	3	3.48	.86
30	4	3.82	.91
35	4	3.62	.90
40	5	3.91	.93
45	5	3.74	.93
50	5	3.60	.92

The best option is n = 45 and d = 5.

It is possible to use other criteria for choosing the pair (n,d). If we feel that the true parametric configuration can in some sense be described by one of two possible slippage configurations given by, say, $p = .75$, $\delta = .05$ and $p = .90$, $\delta = .03$, then we can choose the pair (n,d) that controls the PCS or E(S) at given levels for both configurations.

5. Concluding Remarks

The rule of Gupta and Sobel (1960) which we considered is for the case of samples of equal size. They have also discussed the case of unequal sample sizes. These rules are generally known as unconditional rules. Gupta, Huang and Huang (1976) have discussed conditional rules. The constant used in a conditional rule not only depends on k and n, but also the data $(x_1,\ldots,x_k)$. Finally the processes can be ranked in terms of suitable parametric functions rather than the parameters p_i themselves. A useful ranking function is the entropy function $\theta = p \log p + (1-p)\log(1-p)$. Procedures for selection in terms of θ_i's have been proposed and discussed by Gupta and Huang (`976).

Acknowledgement

The authors are thankful to Professor S. Panchapakesan for help in the preparation of this report and to J. A. Ayres and Joann Ostrowski for programming the relevant computations.

References

Deming, W. E. (1982). Quality, Productivity, and Competitive Position, Center for Advanced Engineering Study, Massachusetts Institute of Technology, Cambridge, Mass.

Gupta, S. S. and Huang, D. Y. (1976). Subset selection procedures for the entropy function associated with the binomial populations. Sankhyā Ser. A 38, 153-173.

Gupta, S. S., Huang, D. Y. and Huang, W. T. (1976). On ranking and selection procedures and tests of homogeneity for binomial populations. In Essays in Probability and Statistics (S. Ikeda et al, eds.), Shinko Tsusho Co. Ltd., Tokyo, Japan, pp. 501-533.

Gupta, S. S. and Sobel, M. (1960). Selecting a subset containing the best of several binomial populations. In Contributions to Probability and Statistics (I. Olkin et al, eds.), Stanford University Press, pp. 224-248.

Table 1A. PCS (top line), P(NCS) (middle line) and E(S) (bottom line)

p= 0.50 delta= 0.10

	k= 3				5				10				15			
n	d=2	3	4	5	2	3	4	5	2	3	4	5	2	3	4	5
	0.95	0.99	1.00	1.00	0.92	0.99	1.00	1.00	0.87	0.97	1.00	1.00	0.83	0.96	1.00	1.00
5	0.87	0.97	1.00	1.00	0.82	0.95	0.99	1.00	0.74	0.93	0.99	1.00	0.70	0.91	0.99	1.00
	2.69	2.93	2.99	3.00	4.19	4.80	4.98	5.00	7.56	9.33	9.92	10.00	10.60	13.69	14.83	15.00
	0.90	0.96	0.99	1.00	0.84	0.93	0.98	0.99	0.75	0.89	0.96	0.99	0.69	0.86	0.95	0.99
10	0.69	0.84	0.93	0.98	0.62	0.79	0.91	0.97	0.51	0.71	0.87	0.95	0.45	0.67	0.84	0.94
	2.29	2.64	2.85	2.95	3.30	4.10	4.61	4.87	5.33	7.32	8.77	9.56	6.98	10.17	12.68	14.14
	0.88	0.94	0.98	0.99	0.81	0.90	0.96	0.98	0.70	0.84	0.92	0.97	0.64	0.79	0.90	0.96
15	0.58	0.73	0.84	0.92	0.50	0.66	0.80	0.90	0.40	0.57	0.73	0.86	0.34	0.51	0.69	0.83
	2.04	2.39	2.66	2.83	2.81	3.56	4.16	4.57	4.26	5.97	7.51	8.66	5.37	7.99	10.51	12.51
	0.87	0.93	0.97	0.98	0.80	0.88	0.94	0.97	0.69	0.81	0.89	0.95	0.62	0.76	0.86	0.93
20	0.50	0.64	0.75	0.85	0.43	0.57	0.70	0.81	0.33	0.47	0.62	0.76	0.27	0.42	0.57	0.72
	1.88	2.20	2.47	2.68	2.50	3.16	3.76	4.23	3.63	5.07	6.51	7.76	4.47	6.62	8.88	10.96
	0.87	0.92	0.96	0.98	0.79	0.87	0.93	0.96	0.68	0.79	0.88	0.93	0.62	0.74	0.84	0.91
25	0.44	0.56	0.68	0.78	0.37	0.50	0.62	0.74	0.28	0.41	0.54	0.67	0.23	0.35	0.49	0.62
	1.75	2.05	2.32	2.54	2.28	2.87	3.43	3.92	3.21	4.45	5.74	6.96	3.88	5.69	7.68	9.66
	0.87	0.92	0.95	0.98	0.79	0.87	0.92	0.96	0.68	0.79	0.86	0.92	0.62	0.73	0.83	0.90
30	0.40	0.50	0.61	0.71	0.33	0.44	0.56	0.67	0.25	0.36	0.48	0.60	0.20	0.31	0.42	0.55
	1.66	1.93	2.18	2.40	2.12	2.64	3.16	3.63	2.91	3.98	5.14	6.30	3.47	5.01	6.77	8.60
	0.87	0.92	0.95	0.97	0.80	0.87	0.92	0.95	0.69	0.78	0.86	0.91	0.62	0.73	0.82	0.88
35	0.36	0.46	0.56	0.65	0.30	0.40	0.50	0.61	0.22	0.32	0.42	0.54	0.18	0.27	0.37	0.49
	1.59	1.83	2.07	2.28	1.99	2.45	2.93	3.39	2.68	3.62	4.66	5.73	3.16	4.50	6.06	7.73
	0.88	0.92	0.95	0.97	0.80	0.87	0.91	0.95	0.70	0.78	0.85	0.91	0.63	0.73	0.81	0.88
40	0.32	0.41	0.51	0.60	0.27	0.36	0.46	0.56	0.20	0.28	0.38	0.48	0.16	0.24	0.33	0.44
	1.52	1.75	1.97	2.17	1.88	2.30	2.74	3.17	2.49	3.33	4.27	5.26	2.92	4.10	5.49	7.01
	0.88	0.92	0.95	0.97	0.81	0.87	0.91	0.95	0.71	0.79	0.85	0.90	0.64	0.73	0.81	0.87
45	0.30	0.38	0.46	0.55	0.25	0.33	0.42	0.51	0.18	0.26	0.34	0.44	0.15	0.22	0.30	0.39
	1.47	1.68	1.88	2.08	1.79	2.18	2.58	2.98	2.34	3.09	3.94	4.85	2.72	3.77	5.02	6.40
	0.89	0.92	0.95	0.97	0.82	0.87	0.91	0.94	0.72	0.79	0.85	0.90	0.65	0.74	0.81	0.87
50	0.27	0.35	0.43	0.51	0.23	0.30	0.38	0.47	0.17	0.23	0.31	0.40	0.14	0.20	0.27	0.36
	1.43	1.61	1.80	1.99	1.72	2.07	2.44	2.81	2.22	2.89	3.67	4.51	2.56	3.50	4.63	5.89

Table 1A (Continued). PCS (top line), P(NCS) (middle line) and E(S) (bottom line)

p = 0.50 delta = 0.10

	k= 3				5				10				15			
n	d=2	3	4	5	2	3	4	5	2	3	4	5	2	3	4	5
	0.90	0.93	0.95	0.97	0.83	0.88	0.92	0.94	0.73	0.80	0.86	0.90	0.67	0.75	0.82	0.87
60	0.23	0.29	0.36	0.44	0.19	0.25	0.32	0.40	0.14	0.20	0.26	0.34	0.12	0.17	0.23	0.30
	1.36	1.51	1.68	1.84	1.60	1.89	2.21	2.54	2.01	2.57	3.22	3.93	2.30	3.07	3.99	5.05
	0.91	0.93	0.95	0.97	0.85	0.89	0.92	0.95	0.76	0.81	0.86	0.90	0.70	0.76	0.82	0.87
70	0.20	0.25	0.31	0.38	0.17	0.22	0.28	0.34	0.12	0.17	0.22	0.29	0.10	0.14	0.19	0.25
	1.30	1.44	1.58	1.72	1.51	1.76	2.02	2.31	1.86	2.33	2.88	3.49	2.11	2.75	3.53	4.43
	0.92	0.94	0.96	0.97	0.86	0.90	0.93	0.95	0.78	0.83	0.87	0.91	0.72	0.78	0.83	0.88
80	0.17	0.22	0.27	0.32	0.14	0.19	0.24	0.29	0.11	0.15	0.19	0.25	0.09	0.12	0.17	0.22
	1.26	1.37	1.49	0.62	1.44	1.65	1.88	2.13	1.74	2.14	2.61	3.13	1.95	2.50	3.16	3.93
	0.93	0.95	0.96	0.97	0.88	0.91	0.93	0.95	0.79	0.84	0.88	0.91	0.74	0.80	0.85	0.88
90	0.15	0.19	0.23	0.28	0.12	0.16	0.21	0.26	0.09	0.13	0.17	0.21	0.08	0.11	0.14	0.19
	1.22	1.32	1.42	1.54	1.38	1.56	1.76	1.97	1.64	1.99	2.39	2.84	1.83	2.30	2.87	3.53
	0.93	0.95	0.96	0.98	0.89	0.92	0.94	0.96	0.81	0.86	0.89	0.92	0.76	0.81	0.86	0.89
100	0.13	0.16	0.20	0.25	0.11	0.14	0.18	0.22	0.08	0.11	0.15	0.19	0.07	0.09	0.13	0.16
	1.19	1.28	1.37	1.47	1.33	1.48	1.65	1.85	1.56	1.86	2.21	2.60	1.73	2.14	2.63	3.20
	0.99	0.99	0.99	0.99	0.98	0.98	0.99	0.99	0.96	0.97	0.97	0.98	0.94	0.95	0.96	0.97
250	0.02	0.02	0.03	0.04	0.02	0.02	0.03	0.04	0.02	0.02	0.03	0.03	0.01	0.02	0.02	0.03
	1.03	1.04	1.05	1.07	1.05	1.07	1.10	1.13	1.10	1.14	1.20	1.26	1.13	1.20	1.28	1.37
	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
500	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	1.00	1.00	1.00	1.00	1.00	1.00	1.01	1.01	1.01	1.01	1.01	1.02	1.01	1.01	1.02	1.02

For values of $n \geq 60$, the values in the above table were computed by using the normal approximations given in (3.3).

Table 1B. PCS (top line), P(NCS) (middle line) and E(S) (bottom line)

p= 0.75 delta= 0.05

	k= 3				5				10				15			
	d=2	3	4	5	2	3	4	5	2	3	4	5	2	3	4	5
n																
	0.97	1.00	1.00	1.00	0.96	1.00	1.00	1.00	0.95	0.99	1.00	1.00	0.94	0.99	1.00	1.00
5	0.94	0.99	1.00	1.00	0.92	0.99	1.00	1.00	0.90	0.99	1.00	1.00	0.90	0.98	1.00	1.00
	2.86	2.98	3.00	3.00	4.65	4.95	5.00	5.00	9.08	9.86	9.99	10.00	13.52	14.78	14.99	15.00
	0.91	0.97	0.99	1.00	0.87	0.96	0.99	1.00	0.80	0.93	0.98	1.00	0.77	0.92	0.98	1.00
10	0.81	0.93	0.98	1.00	0.75	0.90	0.97	0.99	0.68	0.86	0.96	0.99	0.63	0.84	0.95	0.99
	2.54	2.83	2.95	2.99	3.89	4.56	4.87	4.97	6.89	8.69	9.59	9.90	9.63	12.65	14.23	14.82
	0.87	0.95	0.98	0.99	0.81	0.92	0.97	0.99	0.72	0.87	0.95	0.98	0.67	0.84	0.93	0.98
15	0.72	0.85	0.93	0.98	0.64	0.80	0.91	0.97	0.54	0.73	0.87	0.95	0.48	0.69	0.85	0.94
	2.31	2.65	2.85	2.95	3.37	4.13	4.61	4.85	5.57	7.47	8.79	9.52	7.45	10.52	12.79	14.10
	0.85	0.93	0.97	0.99	0.77	0.88	0.95	0.98	0.67	0.82	0.91	0.96	0.61	0.77	0.89	0.95
20	0.65	0.78	0.88	0.94	0.56	0.72	0.84	0.92	0.45	0.63	0.79	0.89	0.40	0.58	0.75	0.87
	2.15	2.49	2.73	2.88	3.02	3.77	4.32	4.68	4.75	6.52	7.98	8.98	6.16	8.93	11.36	13.10
	0.83	0.91	0.96	0.98	0.75	0.86	0.93	0.97	0.63	0.78	0.88	0.94	0.57	0.73	0.85	0.93
25	0.59	0.73	0.83	0.91	0.51	0.66	0.78	0.88	0.40	0.56	0.71	0.83	0.34	0.50	0.66	0.80
	2.02	2.36	2.62	2.79	2.77	3.48	4.06	4.47	4.19	5.80	7.26	8.39	5.31	7.77	10.14	12.07
	0.82	0.90	0.95	0.97	0.73	0.84	0.91	0.95	0.61	0.75	0.85	0.92	0.55	0.70	0.82	0.90
30	0.55	0.68	0.78	0.87	0.46	0.60	0.73	0.83	0.35	0.50	0.64	0.77	0.30	0.44	0.59	0.73
	1.92	2.25	2.51	2.71	2.58	3.25	3.82	4.27	3.79	5.25	6.65	7.84	4.72	6.91	9.13	11.11
	0.82	0.89	0.94	0.97	0.72	0.82	0.90	0.94	0.60	0.73	0.83	0.90	0.53	0.67	0.79	0.88
35	0.51	0.63	0.74	0.83	0.43	0.56	0.68	0.78	0.32	0.45	0.59	0.71	0.27	0.40	0.54	0.67
	1.85	2.16	2.42	2.63	2.43	3.05	3.62	4.08	3.49	4.82	6.15	7.34	4.28	6.24	8.31	10.26
	0.81	0.88	0.93	0.96	0.72	0.81	0.89	0.93	0.59	0.71	0.81	0.89	0.52	0.66	0.77	0.86
40	0.48	0.60	0.70	0.79	0.40	0.52	0.64	0.74	0.30	0.42	0.54	0.67	0.24	0.36	0.49	0.62
	1.78	2.08	2.33	2.55	2.31	2.89	3.44	3.91	3.25	4.47	5.72	6.89	3.94	5.71	7.63	9.51
	0.81	0.88	0.92	0.96	0.72	0.81	0.88	0.93	0.59	0.70	0.80	0.88	0.52	0.64	0.75	0.84
45	0.46	0.56	0.67	0.76	0.37	0.49	0.60	0.70	0.27	0.39	0.51	0.62	0.23	0.33	0.45	0.57
	1.73	2.01	2.26	2.47	2.21	2.76	3.28	3.74	3.06	4.18	5.35	6.49	3.67	5.28	7.07	8.87
	0.81	0.87	0.92	0.95	0.71	0.80	0.87	0.92	0.58	0.70	0.79	0.86	0.51	0.63	0.74	0.83
50	0.43	0.54	0.64	0.73	0.35	0.46	0.57	0.67	0.26	0.36	0.47	0.59	0.21	0.31	0.42	0.53
	1.68	1.94	2.19	2.40	2.13	2.64	3.14	3.60	2.90	3.93	5.04	6.13	3.45	4.93	6.59	8.30

Table 1B (Continued). PCS (top line), P(NCS) (middle line) and E(S) (bottom line)

p = 0.75 delta = 0.05

	k= 3				5				10				15			
n	d=2	3	4	5	2	3	4	5	2	3	4	5	2	3	4	5
	0.81	0.87	0.91	0.94	0.70	0.78	0.85	0.90	0.56	0.66	0.75	0.83	0.48	0.59	0.69	0.78
60	0.39	0.48	0.58	0.66	0.31	0.41	0.50	0.60	0.22	0.30	0.40	0.50	0.18	0.25	0.34	0.44
	1.59	1.83	2.06	2.27	1.96	2.41	2.86	3.30	2.55	3.41	4.35	5.34	2.94	4.11	5.48	6.97
	0.81	0.87	0.91	0.94	0.71	0.78	0.85	0.89	0.57	0.66	0.75	0.82	0.49	0.59	0.68	0.77
70	0.36	0.44	0.53	0.61	0.29	0.37	0.46	0.55	0.20	0.28	0.36	0.45	0.16	0.23	0.31	0.39
	1.53	1.75	1.96	2.16	1.86	2.26	2.68	3.08	2.38	3.14	3.99	4.88	2.72	3.75	4.96	6.30
	0.82	0.87	0.91	0.94	0.71	0.78	0.84	0.89	0.57	0.66	0.74	0.81	0.49	0.59	0.68	0.76
80	0.33	0.41	0.49	0.57	0.27	0.34	0.42	0.50	0.19	0.25	0.33	0.41	0.15	0.21	0.28	0.36
	1.48	1.68	1.88	2.07	1.78	2.15	2.52	2.90	2.25	2.93	3.69	4.51	2.55	3.47	4.55	5.75
	0.82	0.87	0.90	0.93	0.72	0.79	0.84	0.89	0.58	0.67	0.74	0.81	0.50	0.59	0.68	0.75
90	0.31	0.38	0.45	0.53	0.25	0.31	0.39	0.47	0.17	0.23	0.30	0.38	0.14	0.19	0.25	0.33
	1.44	1.63	1.81	1.99	1.71	2.05	2.40	2.75	2.14	2.75	3.45	4.20	2.41	3.24	4.21	5.30
	0.83	0.87	0.91	0.93	0.73	0.79	0.84	0.88	0.59	0.67	0.74	0.80	0.51	0.60	0.68	0.75
100	0.29	0.35	0.42	0.49	0.23	0.29	0.36	0.43	0.16	0.22	0.28	0.35	0.13	0.18	0.23	0.30
	1.41	1.58	1.75	1.92	1.66	1.96	2.28	2.62	2.05	2.61	3.24	3.93	2.30	3.05	3.93	4.93
	0.90	0.92	0.94	0.95	0.84	0.87	0.89	0.91	0.74	0.78	0.82	0.85	0.68	0.72	0.77	0.81
250	0.13	0.15	0.18	0.21	0.11	0.13	0.15	0.18	0.08	0.10	0.12	0.15	0.06	0.08	0.10	0.12
	1.16	1.22	1.29	1.37	1.26	1.38	1.51	1.65	1.44	1.65	1.90	2.16	1.56	1.84	2.17	2.55
	0.97	0.97	0.98	0.98	0.94	0.95	0.96	0.96	0.89	0.91	0.92	0.93	0.85	0.87	0.89	0.91
500	0.04	0.05	0.06	0.07	0.04	0.04	0.05	0.06	0.03	0.04	0.04	0.05	0.02	0.03	0.04	0.04
	1.05	1.07	1.09	1.11	1.08	1.12	1.16	1.20	1.15	1.22	1.30	1.39	1.20	1.30	1.41	1.53

For values of $n \geq 60$, the values in the above table were computed by using the normal approximations given in (3.3).

Table 1C. PCS (top line), P(NCS) (middle line) and E(S) (bottom line)

p= 0.90 delta= 0.03

	k= 3				5				10				15			
	d=2	3	4	5	2	3	4	5	2	3	4	5	2	3	4	5
n																
	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
5	0.99	1.00	1.00	1.00	0.99	1.00	1.00	1.00	0.99	1.00	1.00	1.00	0.99	1.00	1.00	1.00
	2.98	3.00	3.00	3.00	4.96	5.00	5.00	5.00	9.92	10.00	10.00	10 00	14.88	14.99	15.00	15.00
	0.98	1.00	1.00	1.00	0.98	1.00	1.00	1.00	0.97	1.00	1.00	1.00	0.97	1.00	1.00	1.00
10	0.95	0.99	1.00	1.00	0.94	0.99	1.00	1.00	0.93	0.99	1.00	1.00	0.93	0.99	1.00	1.00
	2.88	2.98	3.00	3.00	4.73	4.95	4.99	5.00	9.35	9.88	9.99	10.00	13.99	14.82	14.98	15.00
	0.96	0.99	1.00	1.00	0.94	0.99	1.00	1.00	0.93	0.98	1.00	1.00	0.92	0.98	1.00	1.00
15	0.89	0.97	0.99	1.00	0.86	0.96	0.99	1.00	0.83	0.95	0.99	1.00	0.82	0.95	0.99	1.00
	2.74	2.93	2.98	3.00	4.38	4.82	4.96	4.99	8.39	9.52	9.89	9.98	12.40	14.23	14.82	14.97
	0.94	0.98	1.00	1.00	0.91	0.98	0.99	1.00	0.87	0.96	0.99	1.00	0.86	0.96	0.99	1.00
20	0.83	0.94	0.98	1.00	0.78	0.92	0.97	0.99	0.73	0.89	0.97	0.99	0.70	0.88	0.96	0.99
	2.60	2.86	2.96	2.99	4.05	4.64	4.89	4.97	7.44	8.99	9.68	9.92	10.71	13.28	14.45	14.86
	0.92	0.98	0.99	1.00	0.88	0.96	0.99	1.00	0.83	0.94	0.98	1.00	0.80	0.93	0.98	1.00
25	0.77	0.90	0.96	0.99	0.72	0.87	0.95	0.98	0.65	0.83	0.93	0.98	0.61	0.81	0.92	0.97
	2.47	2.77	2.92	2.98	3.76	4.44	4.79	4.93	6.66	8.41	9.38	9.80	9.35	12.24	13.90	14.64
	0.91	0.97	0.99	1.00	0.86	0.95	0.98	1.00	0.80	0.92	0.97	0.99	0.76	0.90	0.97	0.99
30	0.73	0.86	0.94	0.98	0.66	0.82	0.92	0.97	0.58	0.77	0.89	0.96	0.54	0.74	0.88	0.95
	2.36	2.69	2.87	2.95	3.52	4.24	4.67	4.88	6.04	7.86	9.02	9.62	8.30	11.28	13.26	14.31
	0.90	0.96	0.99	1.00	0.85	0.93	0.98	0.99	0.77	0.90	0.96	0.99	0.73	0.88	0.95	0.98
35	0.68	0.82	0.91	0.96	0.62	0.78	0.89	0.95	0.53	0.72	0.85	0.94	0.48	0.68	0.83	0.92
	2.27	2.61	2.82	2.92	3.31	4.06	4.54	4.80	5.53	7.36	8.65	9.40	7.47	10.42	12.60	13.92
	0.89	0.95	0.98	0.99	0.83	0.92	0.97	0.99	0.75	0.88	0.95	0.98	0.70	0.85	0.94	0.98
40	0.65	0.79	0.89	0.95	0.58	0.74	0.86	0.93	0.49	0.67	0.81	0.91	0.44	0.63	0.79	0.89
	2.19	2.53	2.76	2.89	3.14	3.88	4.40	4.72	5.12	6.90	8.28	9.16	6.81	9.66	11.95	13.47
	0.88	0.95	0.98	0.99	0.82	0.91	0.96	0.99	0.73	0.86	0.94	0.98	0.68	0.83	0.92	0.97
45	0.62	0.76	0.86	0.93	0.54	0.70	0.83	0.91	0.45	0.63	0.77	0.88	0.40	0.58	0.74	0.86
	2.11	2.46	2.70	2.85	2.99	3.72	4.27	4.63	4.77	6.50	7.91	8.89	6.27	8.99	11.32	13.00
	0.88	0.94	0.97	0.99	0.81	0.90	0.96	0.98	0.72	0.85	0.93	0.97	0.67	0.81	0.91	0.96
50	0.59	0.73	0.83	0.91	0.51	0.67	0.80	0.89	0.42	0.59	0.74	0.85	0.37	0.54	0.70	0.83
	2.05	2.39	2.64	2.81	2.86	3.58	4.14	4.53	4.48	6.14	7.57	8.62	5.82	8.41	10.74	12.53

Table 1C (Continued). PCS (top line), P(NCS) (middle line) and E(S) (bottom line)

p = 0.90 delta = 0.03

	k= 3				5				10				15			
n	d=2	3	4	5	2	3	4	5	2	3	4	5	2	3	4	5
	0.86	0.93	0.96	0.98	0.77	0.87	0.93	0.97	0.64	0.78	0.88	0.94	0.56	0.72	0.84	0.92
60	0.53	0.66	0.77	0.86	0.44	0.58	0.71	0.82	0.33	0.47	0.62	0.75	0.27	0.41	0.56	0.70
	1.91	2.24	2.51	2.71	2.54	3.20	3.79	4.25	3.61	5.04	6.46	7.69	4.37	6.47	8.69	10.74
	0.86	0.92	0.96	0.98	0.77	0.86	0.93	0.96	0.64	0.77	0.86	0.93	0.56	0.70	0.82	0.90
70	0.49	0.61	0.73	0.82	0.41	0.54	0.66	0.77	0.30	0.43	0.57	0.69	0.25	0.37	0.51	0.64
	1.83	2.14	2.41	2.62	2.39	3.01	3.58	4.06	3.34	4.63	5.96	7.18	3.99	5.86	7.91	0.90
	0.86	0.92	0.95	0.98	0.77	0.86	0.92	0.96	0.63	0.76	0.85	0.92	0.55	0.69	0.81	0.89
80	0.45	0.57	0.68	0.78	0.38	0.50	0.62	0.73	0.28	0.39	0.52	0.65	0.22	0.34	0.46	0.59
	1.76	2.06	2.32	2.53	2.27	2.84	3.39	3.87	3.12	4.30	5.54	6.73	3.70	5.38	7.26	9.16
	0.86	0.91	0.95	0.97	0.77	0.85	0.91	0.95	0.63	0.75	0.84	0.91	0.55	0.68	0.79	0.88
90	0.42	0.53	0.64	0.74	0.35	0.46	0.58	0.69	0.26	0.36	0.48	0.60	0.21	0.31	0.42	0.55
	1.70	1.98	2.23	2.45	2.17	2.70	3.22	3.70	2.94	4.02	5.18	6.32	3.46	4.99	6.73	8.52
	0.86	0.91	0.95	0.97	0.77	0.85	0.91	0.95	0.64	0.75	0.84	0.90	0.56	0.68	0.79	0.87
100	0.40	0.50	0.61	0.70	0.33	0.43	0.54	0.65	0.24	0.34	0.45	0.56	0.19	0.28	0.39	0.51
	1.65	1.91	2.16	2.38	2.08	2.58	3.08	3.54	2.79	3.78	4.87	5.96	3.27	4.66	6.27	7.97
	0.90	0.93	0.95	0.97	0.84	0.88	0.91	0.94	0.73	0.79	0.84	0.89	0.66	0.73	0.79	0.85
250	0.19	0.24	0.29	0.35	0.16	0.21	0.26	0.32	0.12	0.16	0.21	0.26	0.10	0.13	0.18	0.23
	1.28	1.40	1.53	1.67	1.47	1.70	1.94	2.20	1.80	2.22	2.70	3.24	2.02	2.59	3.26	4.04
	0.96	0.97	0.97	0.98	0.92	0.94	0.95	0.97	0.86	0.89	0.91	0.93	0.81	0.85	0.88	0.91
500	0.07	0.09	0.11	0.13	0.06	0.08	0.10	0.12	0.05	0.07	0.08	0.10	0.04	0.06	0.07	0.09
	1.10	1.15	1.19	1.25	1.18	1.26	1.35	1.45	1.32	1.48	1.66	1.87	1.43	1.65	1.90	2.20

For values of $n \geq 60$, the values in the above table were computed by using the normal approximations given in (3.3).

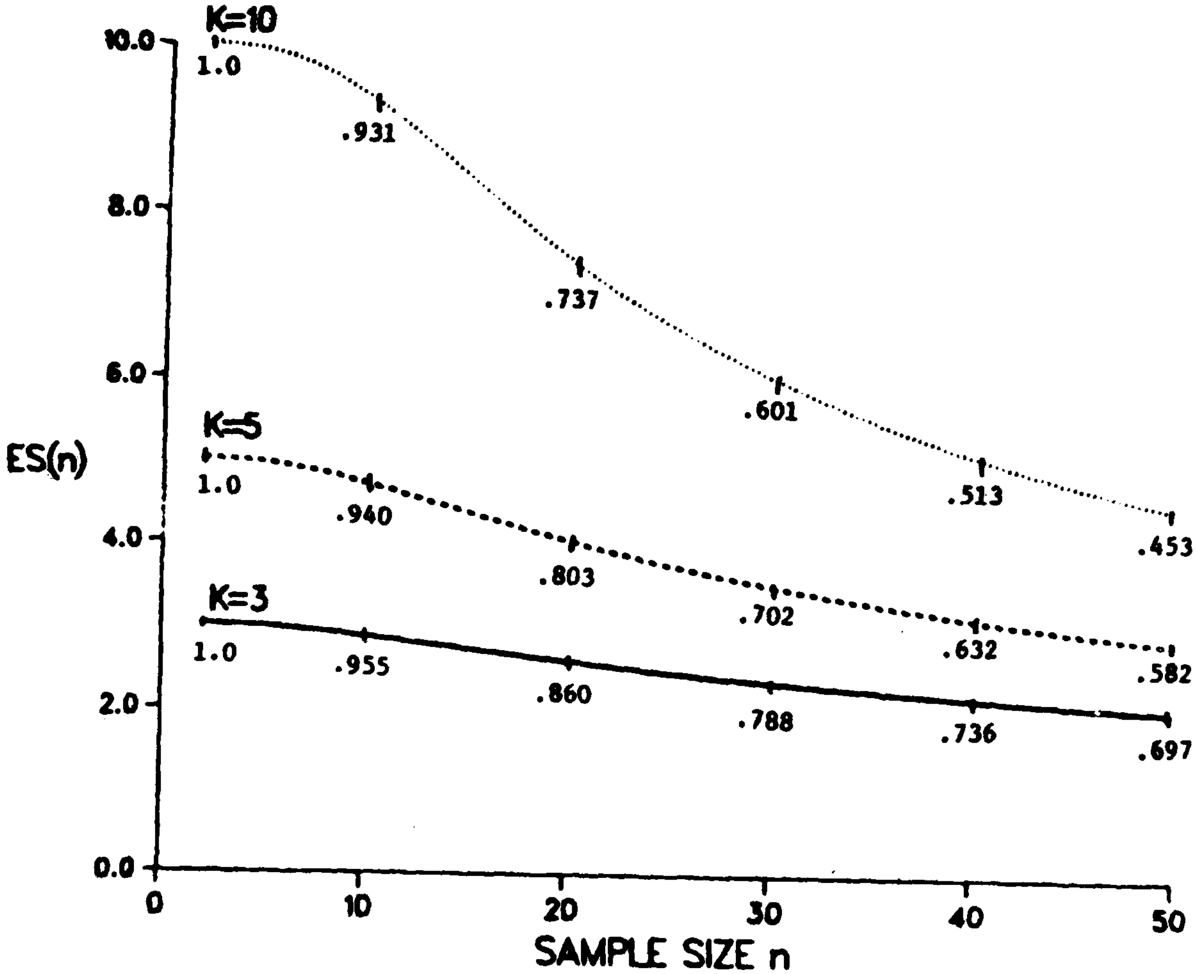
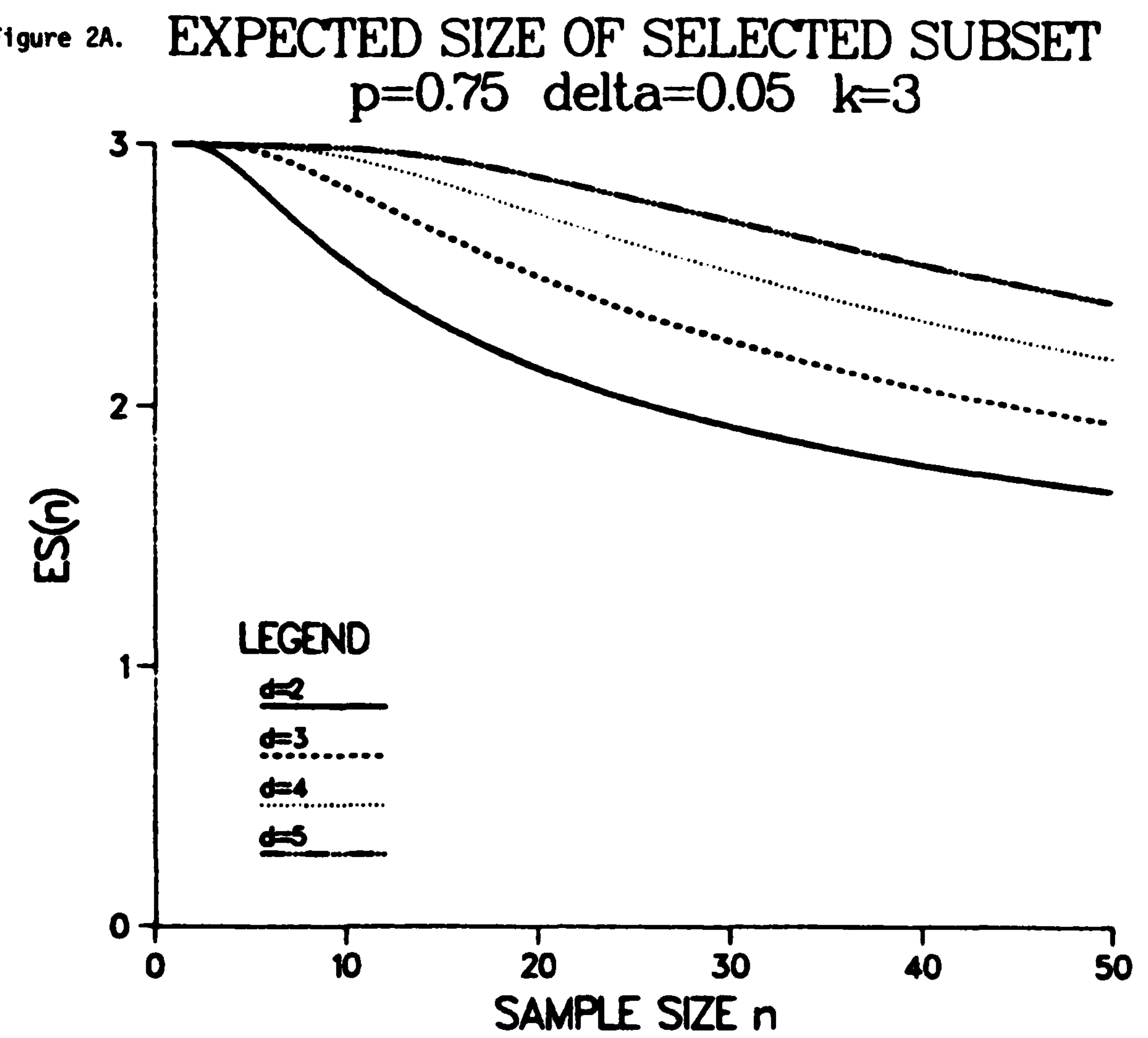


FIGURE 1. Expected size of selected subset for p = .90, δ = .03, d = 2, and k = 3,5,10. Inserted numbers are probability of a correct selection with δ = 0 and p = .90.

Figure 2A.



EXPECTED SIZE OF SELECTED SUBSET
p=0.75 delta=0.05 k=3
ES(n)
3
2
1
0
0
10
20
30
40
50
SAMPLE SIZE n
LEGEND
d=2
d=3
d=4
d=5

Figure 28. EXPECTED SIZE OF SELECTED SUBSET
p=0.75 delta=0.05 k=5

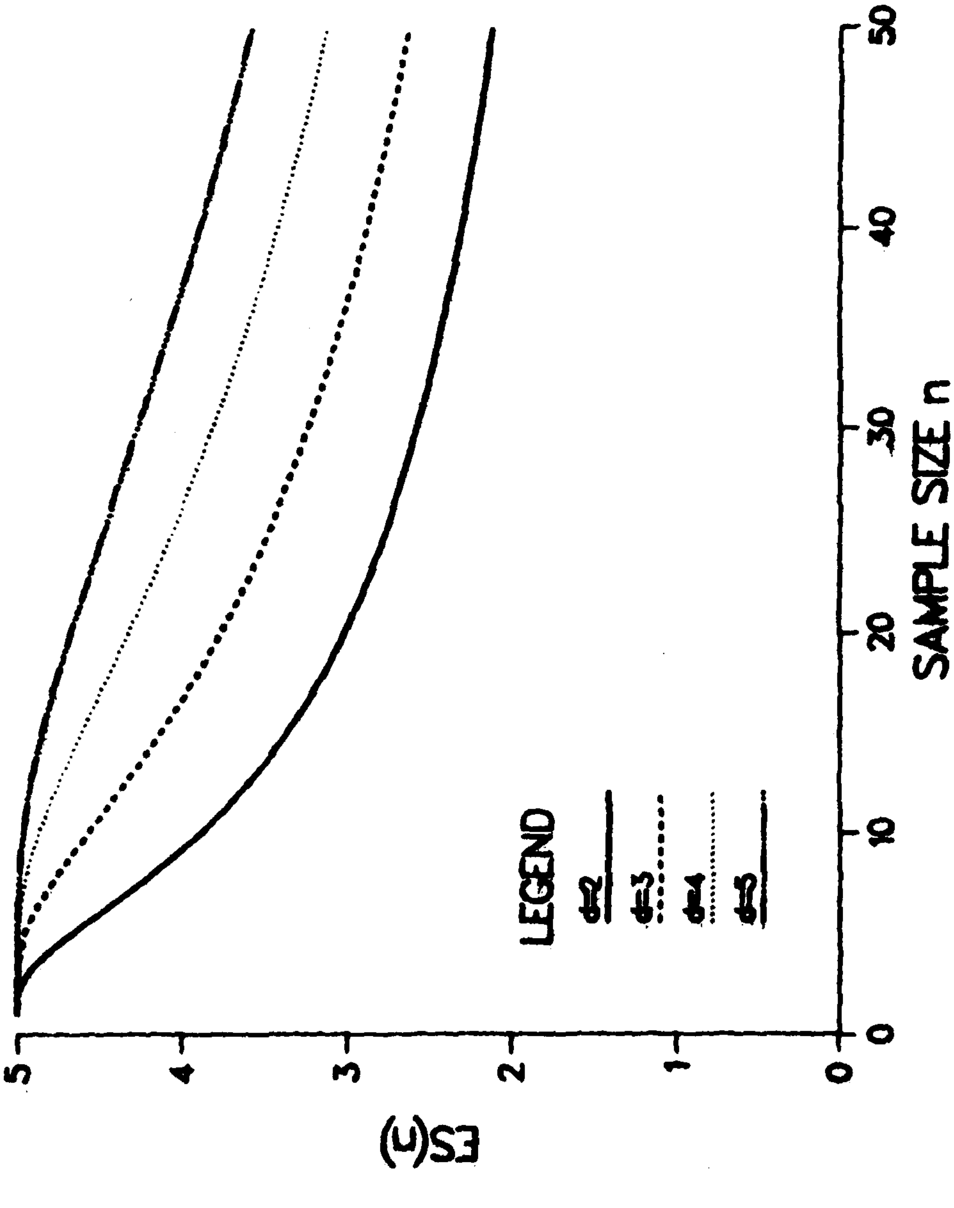


Figure 2C. EXPECTED SIZE OF SELECTED SUBSET
p=0.75 delta=0.05 k=10

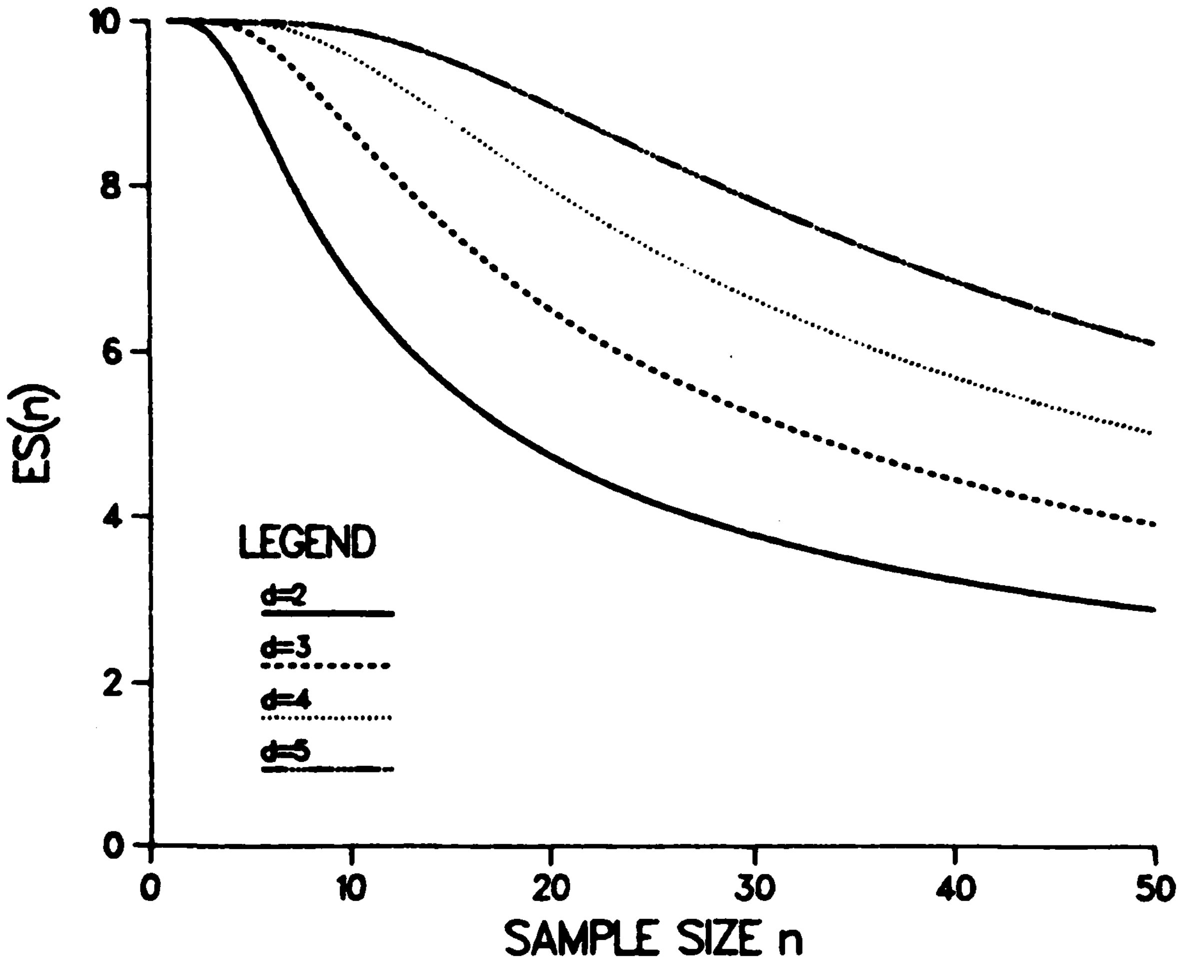
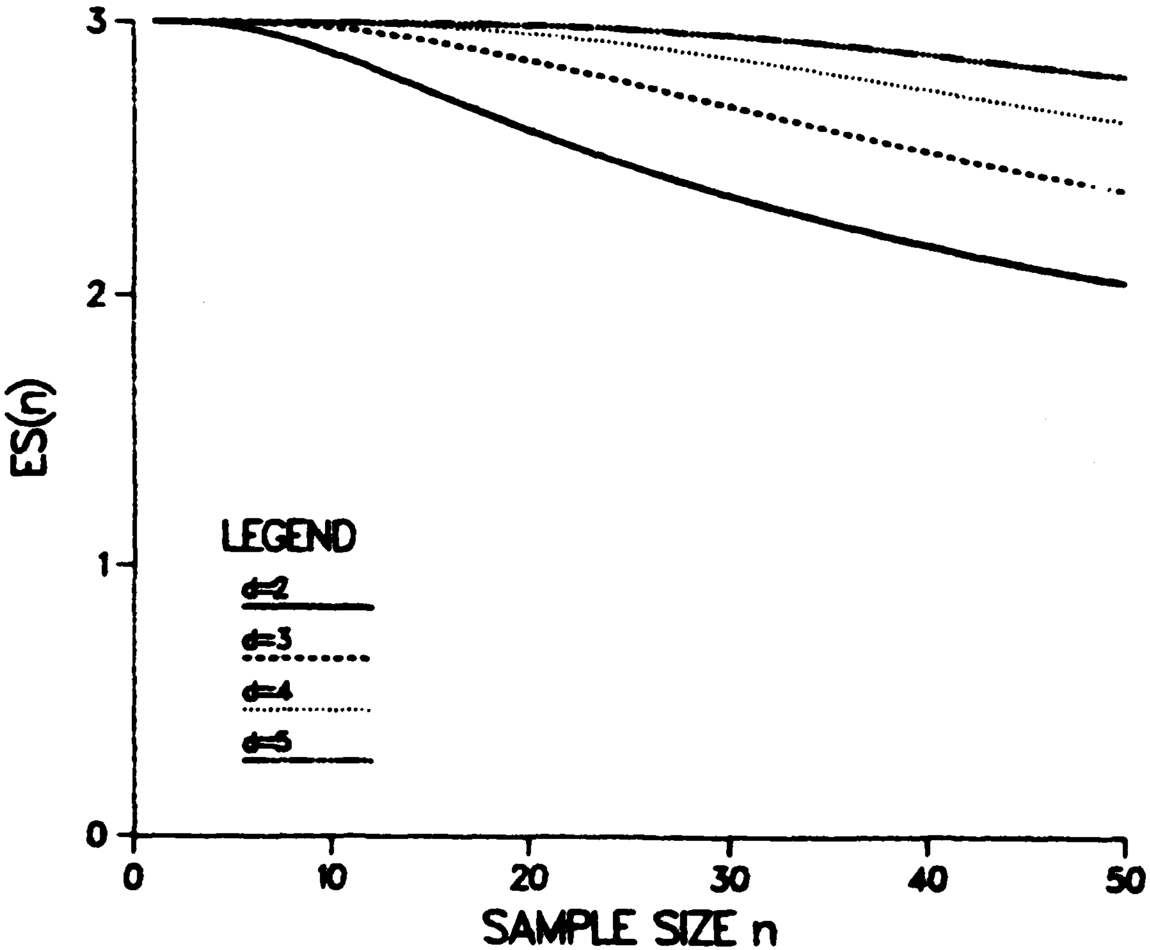


Figure 3A. EXPECTED SIZE OF SELECTED SUBSET
p=0.9 delta=0.03 k=3



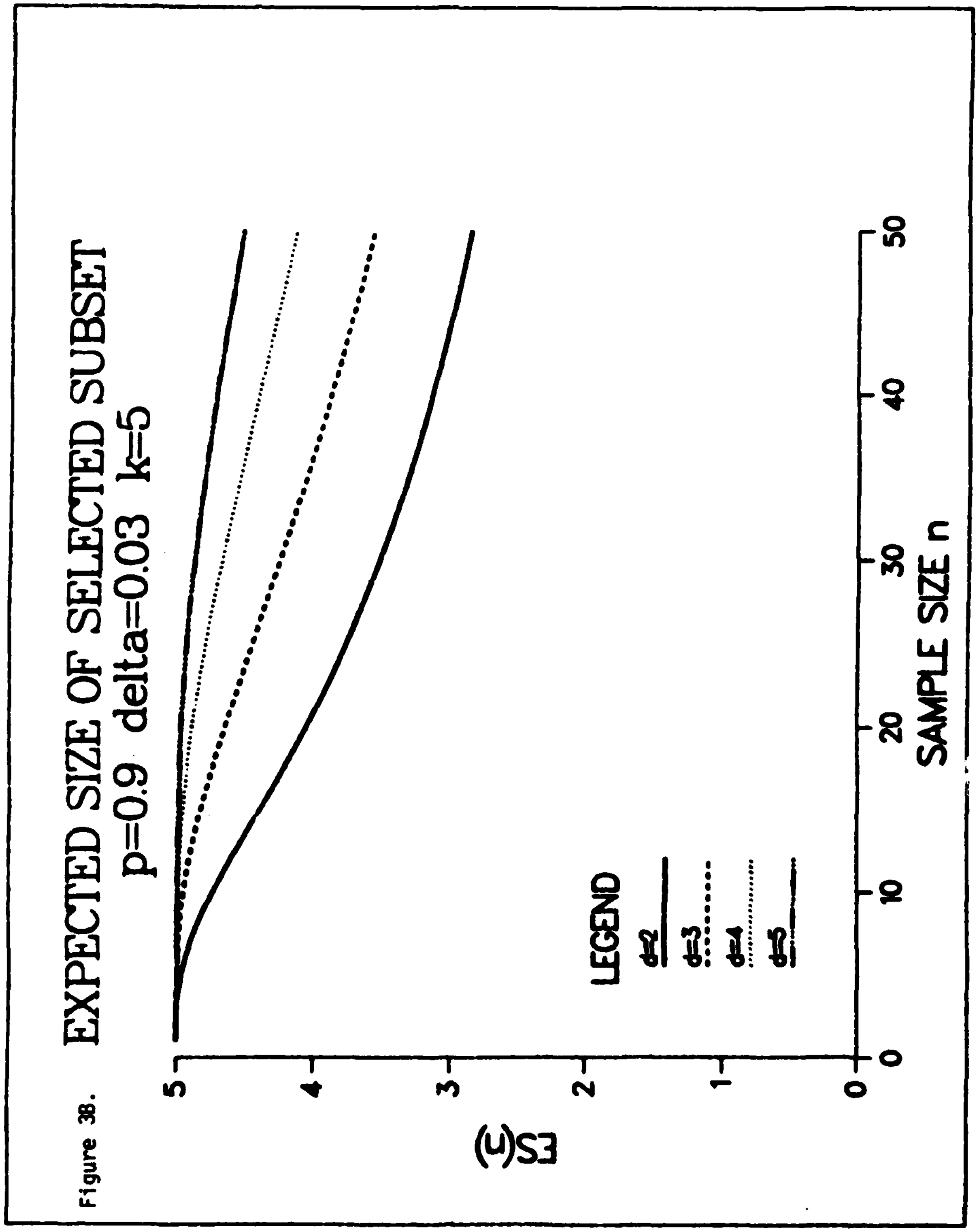
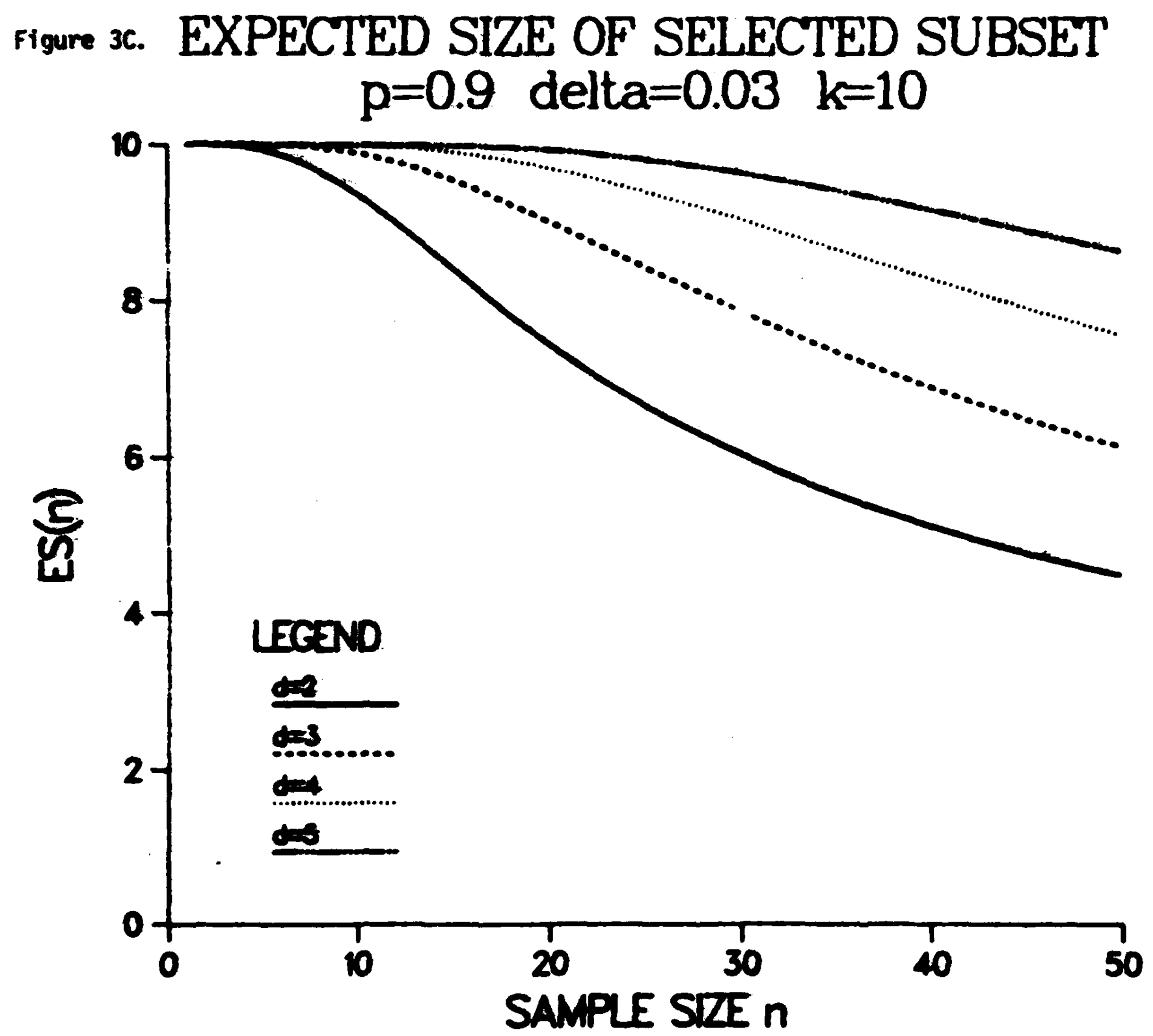

EXPECTED SIZE OF SELECTED SUBSET
p=0.9 delta=0.03 k=5
LEGEND
d=2
d=3
d=4
d=5
ES(n)
5
4
3
2
1
0
0
10
20
30
40
50
SAMPLE SIZE n

Figure 38.

Figure 3C.


EXPECTED SIZE OF SELECTED SUBSET
p=0.9 delta=0.03 k=10
ES(n)
SAMPLE SIZE n
0
2
4
6
8
10
0
10
20
30
40
50
LEGEND
d=2
d=3
d=4
d=5

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER Technical Report #83-44	2. GOVT ACCESSION NO. AD-A135 275	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ON USING SELECTION PROCEDURES WITH BINOMIAL MODELS		5. TYPE OF REPORT & PERIOD COVERED Technical 6. PERFORMING ORG. REPORT NUMBER Technical Report #83-44
7. AUTHOR(s) Shanti S. Gupta and Gary C. McDonald		8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0455
9. PERFORMING ORGANIZATION NAME AND ADDRESS Purdue University Department of Statistics West Lafayette, IN 47907		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Washington, DC		12. REPORT DATE October 1983 13. NUMBER OF PAGES 22
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		15. SECURITY CLASS. (of this report) 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Subset selection rules, operating characteristics, tables, graphs, numerical illustrations, applications to vendor selection.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A subset selection procedure R for binomial populations is considered for the problem of selecting the best of k vendors whose manufacturing processes have the probabilities $p_1,\dots,p_k$ of turning out an item which conforms to specications. Let $X_1,\dots,X_k$ denote the number of conforming items from samples of size n from the k processes. Then the rule R is of the form: Select π_i if and only if $X_i \geq \max_{1\leq j\leq k} X_j - d$, where d is a nonnegative integer. The operating characteristics (over)

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

cont

(i.e. selection probabilities and expected size of the selected subset) of this rule are related to the underlying p_i's, the common sample size n, and d. Formulae (both exact and asymptotic) are given for these quantities for slippage as well as equi-spaced parametric configurations. Tables and graphs relating these quantities are presented for three specific slippage configurations. Numerical illustrations are given to show the use of the tables in determining the sample size n and the constant d to be used in the rule R.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)